Spring 5-1-2016

Computational Investigations into the Molecular Underpinnings of Eyesight Signaling Pathways

Shaan Kamal
University of Connecticut, shaan5kamal@gmail.com

Computational Investigations into the Molecular Underpinnings of Eyesight Signaling Pathways

University Scholar Project

Shaan Kamal

Advisor: Dr. Eric May

Table of Contents

Abstract

Phosphodiesterase 6 (PDE6) is a critical enzyme in the eyesight-signaling pathway. When activated, PDE6 hydrolyzes cGMP to GMP, which deactivates cGMP-gated ion channels, causing hyperpolarization of the cell and activating the sensory neurons responsible for vision. Within the PDE family, PDE6 is the only enzyme known to have an inhibitory subunit (PDE6-γ), which allows for the regulation of cGMP levels. When PDE6-γ is bound to PDE6, the enzyme is turned "off" and cannot catalyze cGMP. The α subunit of the G-protein transducin removes PDE6-γ and activates PDE6. PDE6 has proven problematic to isolate, making it difficult to study experimentally and preventing a structure from being solved. A chimera of the homologous proteins PDE5 and PDE6 (PDE5/6) has been constructed to serve as a model for PDE6 that can be expressed and purified. The validity of PDE5/6 as a model for PDE6 has been demonstrated through mild inhibition by PDE6-γ. To study the native sequence of PDE6, a computational approach was applied by creating a homology model of PDE6. Using this model has allowed us to understand the structural basis for PDE6's inhibition by PDE6-γ.

We have investigated correlations in protein dynamics possibly responsible for allosteric properties by running over 1 microsecond (μs) of molecular dynamics simulations and carrying out various analyses. PDE6 and PDE5/6 are both inhibited by the PDE5 inhibitor sildenafil. With this knowledge, systems of the homology model of PDE6, PDE5, and PDE5/6 in ligand-free (apo), sildenafil bound, and PDE6-γ bound states were constructed and simulated for 100-300 nanoseconds each. Our analyses were aimed at evaluating the quality of our homology model of PDE6 and showing the affect

of sildenafil and PDE6-γ within each system and across different systems. These studies have allowed us to gain insight into the structural characteristics and the specific residues that are involved in allosteric communication and regulation of PDE6's inhibition by PDE6-γ. Using RMSF and cross-correlational analyses we have discovered a novel region that shows correlation with the H- and M-loops and could be an allosteric therapeutic target for diseases in which PDE6 is implicated, such as retinitis pigmentosa. Through this work, we hope to gain a greater understanding of the molecular dynamics and mechanisms of PDE6 and how these features may relate to the molecular underpinnings of eyesight.

Introduction

In Madeleine L'Engle's *A Wrinkle in Time*, the main character attempts to explain vision to a species with no history of eyesight. After the conversation, the character realizes that her perception of the world is so influenced by eyesight that she can't even describe the sense. Just how is it that we see? The foundation of eyesight is a series of molecular level interactions. An integral protein in the molecular pathway that allows us to experience eyesight is phosphodiesterase 6 (PDE6). Mutations and structural changes in PDE6 have been implicated in diseases of the eye such as retinitis pigmentosa (RP) and congenital night blindness. Despite the significance of PDE6 in the eyesight signaling pathway and its relationship to diseases of the eye, it has not been very well studied.

The goal of this project was to elucidate the role of structural dynamics in the inhibition/activation mechanisms of PDE6 through a computational investigation. A powerful aspect of computational biology is the use of molecular dynamics methods to simulate molecular interactions and gain insight in ways not possible through experimental techniques. These methods effectively allow us to use computers as "molecular microscopes" to view molecular level interactions[1]. Although the groundwork for these techniques was laid down decades ago, computer technology has just recently progressed far enough to meet the technical specifications required of intensive simulations, making computational biology a very exciting and emerging field. Due to the challenges in studying PDE6 experimentally, a computational approach has the potential to greatly increase our knowledge of this important enzyme.

The aim of this study was to gain a thorough understanding of the structure-function relationships and critical residues in PDE6 by carrying out molecular dynamics simulations, dynamic cross-correlational analysis, homology modeling, and principal component analysis on the enzyme. This knowledge will provide new insights into PDE6 function, which may help to elucidate the cause of RP and congenital night blindness as well as give us a broader understanding of the processes that allow vision to occur.

Review of Literature

To propagate signals, cells use multi step, and often complex, signaling pathways. Cyclic AMP (cAMP) and cyclic GMP (cGMP) are two signaling molecules that help propagate the signals of the pathways in which they are involved. The main function of the class of enzymes called phosphodiesterases (PDEs) is to hydrolyze cAMP and cGMP to AMP and GMP respectively[2]. Given that regulation of cAMP and cGMP allows the various enzymes downstream of cAMP and cGMP in the signaling cascade to also be regulated, PDEs are an important class of regulatory enzymes. Due to their important functions, they are often targeted for drug development[3,4].

PDE6 is one of the eleven different types of PDEs and is found in both the rod and cone cells of the eye[2,5]. One of the properties of PDE6 that makes it unique amongst PDEs is that it has a naturally occurring inhibitory gamma (γ) subunit[6,7]. When the γ-subunit is bound to PDE6, the enzyme is inactive and when removed PDE6 becomes active[6,7]. This allows the γ-subunit to regulate the activity of PDE6. Although there are several mechanisms by which PDEs can be inhibited, the γ-subunit is known to affect only PDE6[6]. Two structural features, the H-loop and the M-loop, near the binding site of

cGMP are thought to significantly aid the γ-subunit in binding PDE6[8,9]. PDE6 has multiple domains and we are only studying the catalytic domain (the catalytic domain of PDE6 alone has been shown to be functional and can be inhibited by the γ-subunit)[8, 10]. The catalytic domain is a dimer of two identical subunits (αα) in cone PDE6, and the focus of this study will be the catalytic domain monomer (α)[10].

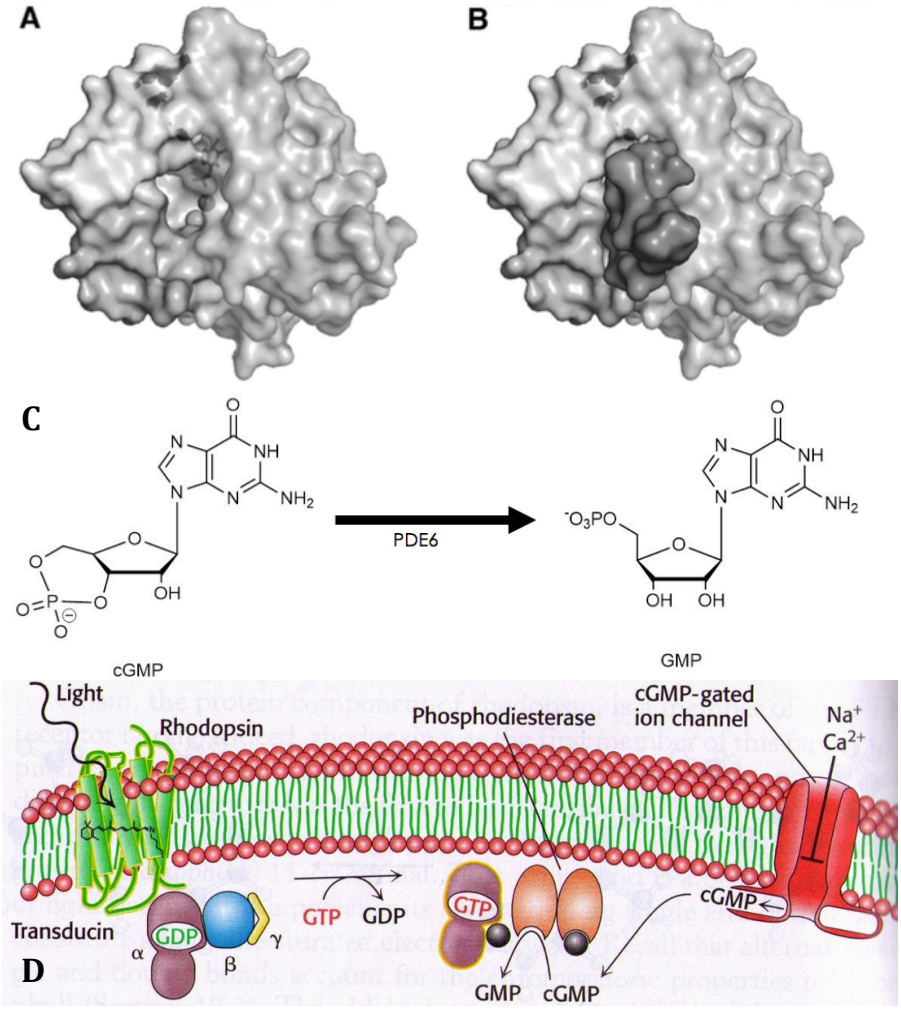


Figure 1: A) Surface representation of PDE6 model showing the catalytic pocket[8]. B) Surface representation of PDE6 model highlighting the γ-subunit (dark grey) blocking access to the catalytic pocket of PDE6. C) The reaction of cGMP to GMP catalyzed by PDE6. D) Signaling Pathway of Eyesight Diagram[11]

The signaling pathway of eyesight has several steps, as shown in Figure 1D. When light enters the eye, it interacts with the G-protein coupled receptor rhodophsin,

which in turn activates the G-protein transducin[12]. The α-subunit of transducin then activates PDE6 by interacting with the γ-subunit and dissociating it from PDE6, allowing PDE6 to rapidly hydrolyze cGMP to GMP[12]. The drop in concentration of cGMP deactivates cGMP-gated ion channels, causing hyperpolarization of the cell. This activates the sensory neurons responsible for vision[6, 12]. Mutations to residue ASN661 on the H-loop (N661S and N661A) of PDE6 have been identified in cases of RP, indicating this residue is critical for proper functioning[8, 13].

Although PDEs are well-studied enzymes, PDE6 has not been as well characterized as other PDEs due to the fact that it is extremely difficult to express in bacteria. As a result, it has never been isolated and its atomic structure has not been solved[6]. To work around this issue, researchers have created a 'chimera' of PDE5 and PDE6, called PDE5/6. PDE5 and PDE6 are very sequentially homologous, so by substituting the sequences unique to PDE6 for the same areas on the PDE5 gene, a hybrid of PDE5 and PDE6 (PDE5/6) can be expressed and isolated[2,6,8]. The PDE5/6 has been shown to retain the ability to hydrolyze cGMP while also being effectively inhibited by the γ-subunit, making it a reasonable model to gain insight into PDE6[8]. The functionality of PDE5 and PDE5/6 differs from that of PDE6 in that the former hydrolyze cGMP with low efficiency (~0.55 cGMP/second) while the latter hydrolyzes cGMP very efficiently (~2000 cGMP/second)[8]. The structure of the PDE5/6 bound with the inhibitory drug sildenafil and the γ-subunit has been determined by X-ray crystallography, providing a basis for molecular modeling studies[8].

Despite the advances brought about by the chimera studies, there are several questions about PDE6 that remain unanswered. The chimera has been shown to have an

H-loop 26 angstroms closer to the M-loop than in PDE5, even though the H-loop amino acid sequence is exactly the same in PDE5 and the chimera[8]. This observation has raised several questions, including: What structural characteristic causes this movement of the H-loop in the chimera? Is the same H-loop configuration present in native PDE6 or is it only observed for the chimera? What structural influence does the γ-subunit have on PDE6? Are there any allosteric interactions occurring in PDE6? What are the structural differences between drug bound, γ-subunit bound, and free states of PDE6? These questions are well suited for the methods of computational biophysics, which I aim to apply in this project.

Methods

Molecular dynamics was first applied to the study of proteins in the 1970s[1, 14]. Using experimental data from physical chemistry, 'force fields' that describe the bonded (i.e.- chemical bonds, bond angles, and bond dihedrals) and non-bonded (i.e.- van der Waals forces and electrostatics) interactions between atoms were developed. In a molecular dynamics simulation, the potential energy at every step is calculated using these force fields for all of the atoms within the system. The software used will then take the negative gradient of these potential energies to get the force acting on every atom. At each time point of the simulation, the atoms are displaced along their force vector to generate a new system configuration. This is repeated thousands of times to get a trajectory of a desired length. The application of molecular dynamics allows us to study proteins at the molecular level in a way not possible with traditional experimental methods.

To carry out my research, I used a computer running a Unix based operating system with Gromacs software installed, and access to a High Performance Computing (HPC) cluster. Access to a computer used for post simulation analysis was provided through the May Lab. The BECAT center provides access to its HPC cluster (HORNET), which was used to carry out simulations that required large amounts of memory and processing power. The local computer and HORNET are optimized to run computationally intensive processes. GROMACS software was used to run the simulations with the CHARMM27 force field and TIP3P water[15].

In order to simulate a protein using molecular dynamics, initial Cartesian coordinates are needed. Most commonly, a structure file is acquired from the Protein DataBase (pdb) where experimentally determined protein structures are deposited. In order to study the PDE5/6 chimera, the PDE5/6 with the γ-subunit bound structure and PDE5/6 with inhibitor sildenafil bound structure (pdb files 3JWR and 3JWQ respectively) were selected for study. The first step taken was to remove the sildenafil from the 3JWQ structure to get an apo PDE5/6 structure. The PDE5 structures used were apo PDE5 and sildenafil bound PDE5 (pdb codes 2H40 and 2H42 respectively). A homology model of PDE6 was then built with the program I-Tasser[16]. The sequence for PDE6 was fed into the I-Tasser server, which then analyzed the sequence of PDE6, found sequentially similar proteins, analyzed the structures of those proteins, and then used all of the structural information to create a homology model of PDE6. Using the modeling and visualization program Chimera, sildenafil and γ-bound PDE6 structures were created[17].

All three of the PDE5/6 structures (apo, sildenafil bound, and γ-bound), the two PDE5 structures (apo and sildenafil bound), and the PDE6 structures (apo, sildenafil bound, γ-bound, N661S mutation, and N661A mutation) were simulated for 100-nanoseconds (Figure 2). The purpose of these simulation is to allow the structures to first reach a "relaxed" conformation, after which the equilibrium dynamics of the enzymes can be analyzed[18].

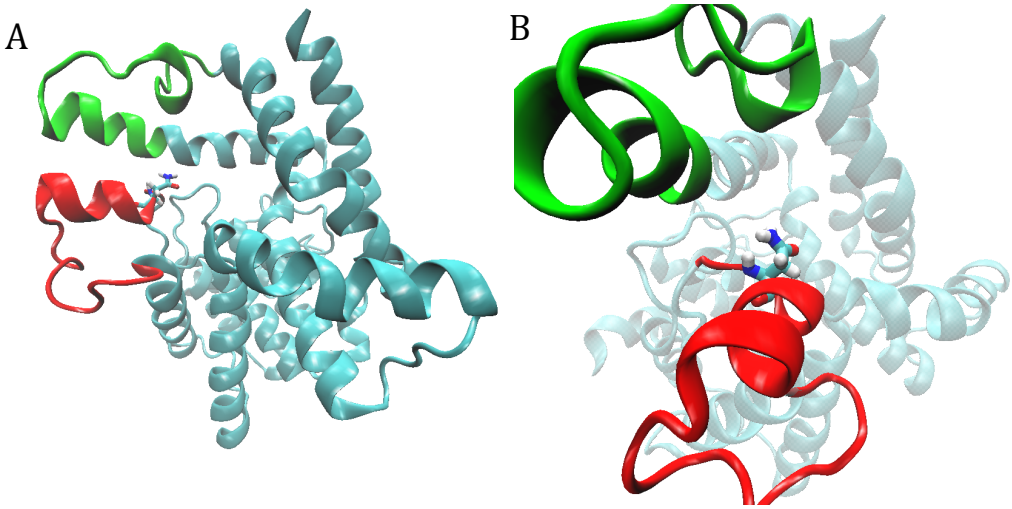


Figure 2: A.) PDE5/6 after 20 nanoseconds of the 100 nanosecond simulation, with the M-loop shown in green and the H-loop shown in red. B.) Residue ASN661 highlighted on the H-loop

The main analyses performed on the 100ns simulation data for each system were root mean square deviation (RMSD), root mean square fluctuation (RMSF), calculation of H- and M-loop distances, principal component analysis (PCA), and cross correlational analysis using the Bio3D suite[19]. These methods were used for validation, system comparisons, and gaining insight into structural dynamics of the proteins.

RMSD is used to calculate how far from the starting structure the proteins' structures deviate over the course of the simulation. RMSD was one of the tools used to assess validity of the PDE6 homology model. The structural deviations of the model measured by RMSD values being equal to or less than the chimera's structural deviations

would verify that the model can be used as an accurate representation of PDE6 for our purposes. RMSF is used to get the fluctuations of each individual amino acid in each protein over the course of the simulations. The overall RMSF data can be used to understand which regions of the protein are the most or least flexible, giving insight into structural dynamics.

The H- and M-loop distances were calculated by measuring the distance between the center of masses (COM) between the H- and M-loops as well as the distances between the residues on the ends of each loop (residues 668 of the H-loop and 797 of the M-loop). Results between the COM and two residue distances were compared to see if the two methods gave agreeable distances and to see if one method is more accurate than the other.

The PCA and cross-correlational analysis are both aimed at getting information about the global, correlated motions of the proteins. PCA involves building a covariance matrix from the trajectory of each simulation, diagonalizing the matrix to get eigenvectors and eigenvalues, and then sorting the eigenvectors by eigenvalue to find the principal components. Each principal component accounts for a percentage of the total positional fluctuations, or global motions, of the proteins.

Principal component analysis (PCA) will be the final tool used in piecing together a greater understanding of the H- and M-loop interface. By taking the structures of PDE5 (pdb file 2H40), the PDE5/6 chimera, mutated PDE6 and the PDE6 homology model and comparing them with principal component analysis, correlated global motions throughout the proteins can be characterized[20]. This will help to elucidate the allosteric interactions occurring within PDE6 that allow it, but not PDE5, to bind the γ-subunit. The mutated

system will give us insight into what structural interactions cause decreased binding of the γ-subunit and the implications in RP.

Collectively, these experiments will allow us to gain insight into the structural basis for inhibition of PDE6, what structural effect mutations have, which conformations of PDE6 are most stable, and what specific interactions within PDE6 allow inhibition by the γ-subunit. The knowledge gleaned through my University Scholar project will add to our knowledge of eyesight signaling pathways and help to understand how diseases such as retinitis pigmentosa occur.

Results

Validation of Methods and Systems

System	Simulation Time	Max RMSD	Max H- and M-loop Distance
PDE5 Apo	100 nanoseconds	3.7 Å	32.6 Å
PDE5 with Sildenafil Bound	100 nanoseconds	2.7 Å	31.0 Å
PDE5/6 Apo	100 nanoseconds	1.8 Å	18.1 Å
PDE5/6 with Sildenafil Bound	100 nanoseconds	2.0 Å	15.7 Å
PDE5/6 with γ-Subunit Bound	100 nanoseconds	2.3 Å	15.9 Å

PDE6 Apo	100 nanoseconds	3.8 Å	12.2 Å
PDE6 with Sildenafil Bound	100 nanoseconds	3.2 Å	12.4 Å
PDE6 with γ-Subunit Bound	100 nanoseconds	3.2 Å	10.9 Å

Table 1: The maximum RMSD and H- and M-loop distance for each of the systems simulated.

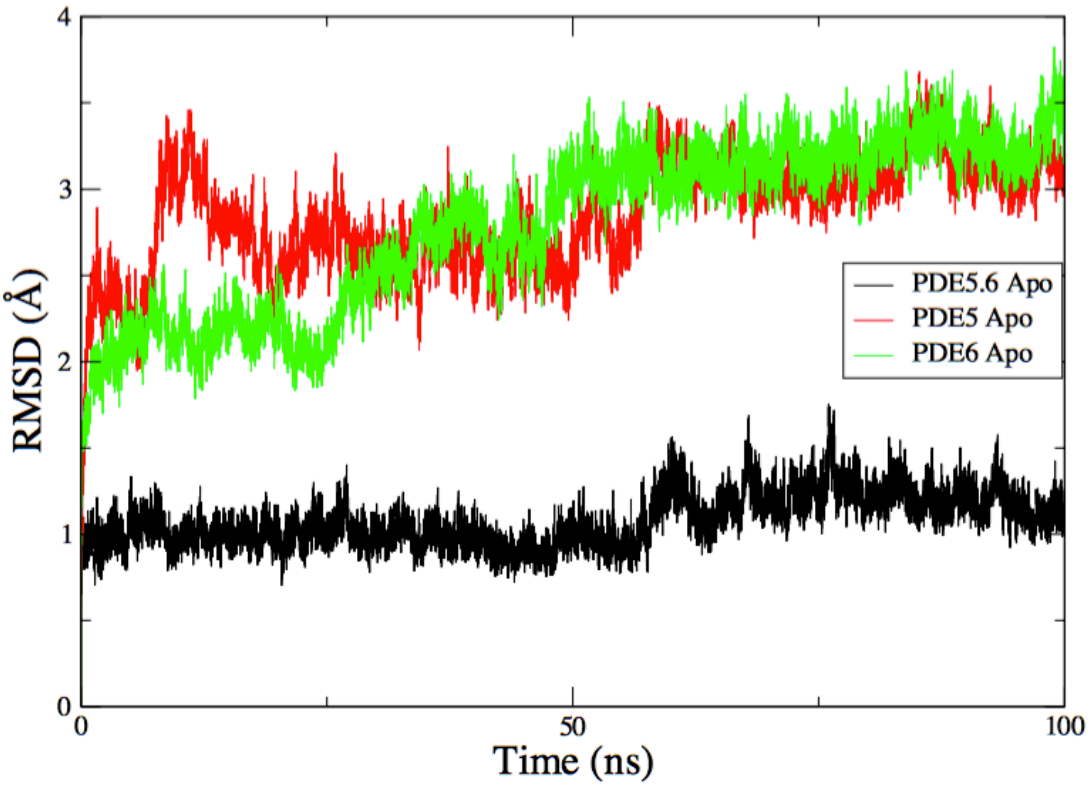


Figure 3: The RMSDs for the PDE5/6, PDE5, and PDE6 apo systems over the course of 100 nanosecond simulations

The initial analyses were done in order to validate the methods and systems used. The RMSD analysis of all of the systems was very low, with the apo systems (Figure 3) showing a maximum RMSD value of only ~4 Å, a low RMSD. A key point to note here

is that the RMSD of the homology model for PDE6 also matches up well with the RMSDs of the real crystal structures. This is a strong indication that the model used is a stable model compatible with our methods and that it is likely an accurate model of PDE6.

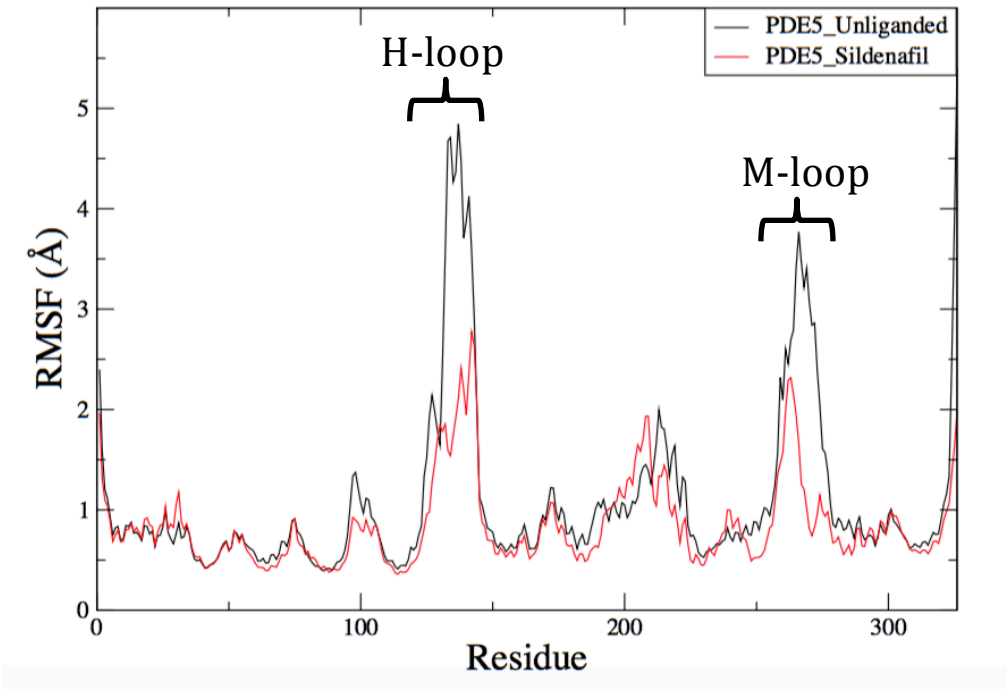


Figure 4: RMSF of PDE5 apo and sildenafil systems after 100 nanosecond simulations. H- and M-loop regions are highlighted.

Because PDE5 is so well studied, it was logical to see if analyses of the simulations gave results that match up well with the knowledge gained from experimental studies. Sildenafil is a known potent inhibitor of PDE5 and therefore the expectation is that the analyses run on the PDE5 apo and sildenafil bound simulations would show that sildenafil causes decreased flexibility, indicating it is a tightly bound inhibitor and would prevent conversion of cGMP. The RMSF data for the PDE5 apo and sildenafil bound simulations (Figure 4) shows that the H- and M-loops of PDE5 apo show very high flexibility, matching what previous experimental studies showed[21]. The RMSF of

sildenafil bound PDE5 shows very reduced flexibility in the H- (residues 126 to 149) and M- (residues 254 to 277) loops, with the RMSFs going from ~5 Å to ~2.5 Å for the H-loop and ~3.5 Å to ~2 Å for the M-loop. These large reductions in the flexibility of these regions is in line with the idea that a potent inhibitor such as sildenafil reduces enzyme flexibility and could give insight into the long range effects of sildenafil binding. The PCA of the PDE5 apo and sildenafil bound simulations allows for further investigation. The apo simulation clearly shows that the system samples a large conformational space. The sildenafil bound simulation shows that the motions of the protein are severely restricted in comparison, with the system only sampling a small portion that the apo system samples. Taken together, these two analyses are very strong validations of the methods we used. The sildenafil bound PDE5 simulation shows reduced flexibility in highly flexible regions (the H- and M-loops) and reduced motions overall of the system. Reduced flexibility and motion are indicative of strong inhibition, which was expected.

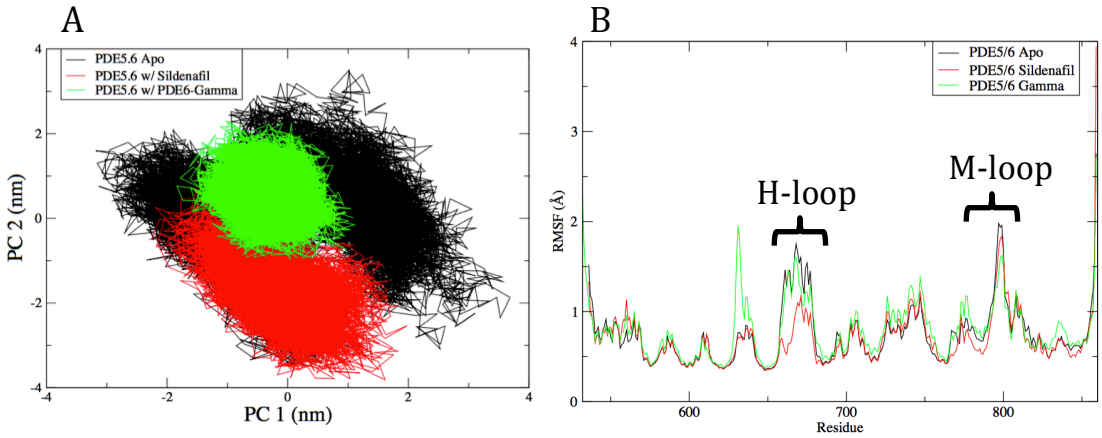


Figure 5: A) Principal component analysis of PDE5/6 apo, PDE5/6 with sildenafil bound, and PDE5/6 with the PDE6 γ-subunit bound. B) RMSF of the PDE5/6 systems after 100 nanosecond simulations

Further validation was done through PCA of data of all of the PDE5/6 simulations (Figure 5a). The apo system shows a large conformational space being sampled, whereas

the sildenafil bound system samples a smaller portion of the conformational space sampled by the apo system. This shows, just as was the case with PDE5, that the sildenafil reduces the overall motions of the protein as is expected of an inhibitor. Interestingly, the γ-bound system showed very little motion, but sampled a space not explored as much as in the apo or sildenafil bound systems. The RMSF data backs this up, as the fluctuations in some regions for the γ-bound system are higher than in the apo system (Figure 5b).

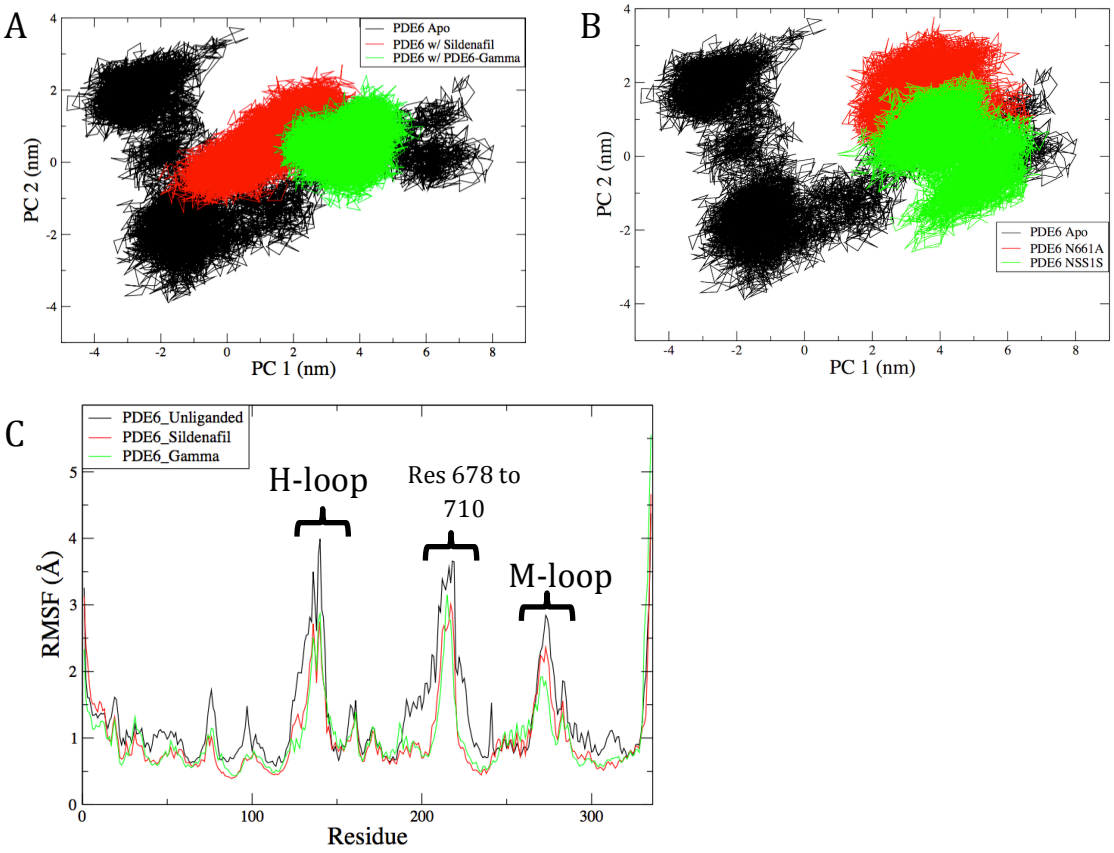


Figure 6: A) Principal component analysis of PDE6 apo, PDE6 with sildenafil bound, and PDE6 with γ-subunit bound. B) Principal analysis of PDE6 apo, PDE6 N661A, and PDE6 N661S. C) RMSF of PDE6 apo, PDE6 with sildenafil bound, and PDE6 with γ-subunit bound

All of this analysis is good validation of the methods used and allowed us to proceed with analyzing the data from the PDE6 systems (Figure 6). The PCA of the apo system shows large sampling, whereas the sildenafil bound system samples a smaller and

different area of the principal component space (Figure 6A). The γ-bound subunit shows a very small sampling of the same PC space that the apo system samples (Figure 6A). Looking at the mutations, the PCA shows that both significantly change the motions of PDE6 in the PC space (Figure 6B). That single point mutations so significantly change the motions of PDE6 is a reflection of the importance of that residue as well as of those mutations in the development of retinitis pigmentosa.

H- and M-Loop Distances and Dynamics

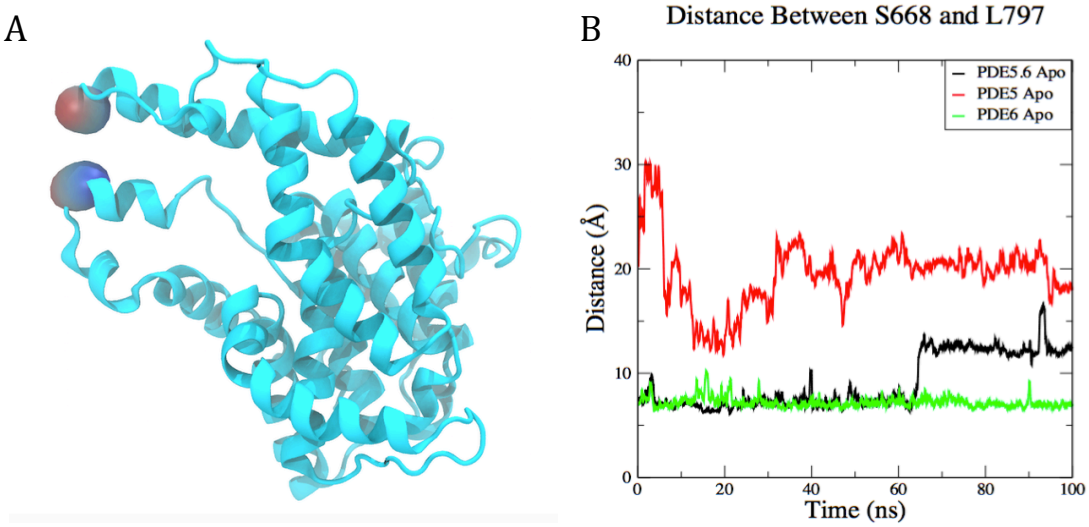


Figure 7: a) H- and M-loop distances were calculated as the distance between residues S668 and L797 (highlighted in the figure) for each system b) H- and M-loop distances for the apo systems.

The H- and M-loop distances were calculated as the distance between the ends of the two loops (Figure 7). Over the course of the apo simulations, PDE5 showed variable distance between the loops, with the distance ranging from ~12 to ~30 Å. PDE5/6 showed a constant distance of ~7 Å before shifting up to ~12 Å. PDE6 meanwhile showed a constant distance of ~7 Å. All of the other PDE6 systems also showed a constant distance of ~7 Å.

The variable PDE5 distance may reflect an equilibration process from the crystal structure or more interestingly may be highlighting that the loops are very dynamic. The constant PDE6 distance mirrors the experimental data from PDE5/6 and could indicate that PDE5/6's H- and M-loop distance may be representative of PDE6's.

The RMSFs for the apo systems highlighted that the PDE5 and PDE6 H- and M-loops were very flexible, indicating that they are dynamic (Figures 4 and 6C). The RMSF of apo PDE5/6 however showed very low flexibility in the residues that make up the H-loop as well as lower flexibility in the M-loop as compared to in PDE5 and PDE6 (Figure 5B). Also, the RMSFs of PDE6 show a third region of high flexibility between the H- and M-loops (residues 678 to 710). This new region showing high flexibility may indicate it is important in the dynamics of PDE6. Interestingly, although the RMSF in the H- and M-loops in PDE6 is high, the distance between the loops is still constant. A reasonable conclusion to reconcile these two data is that the motions of the two loops are correlated, which can be tested by the cross-correlational analysis.

Cross-Correlational Analysis

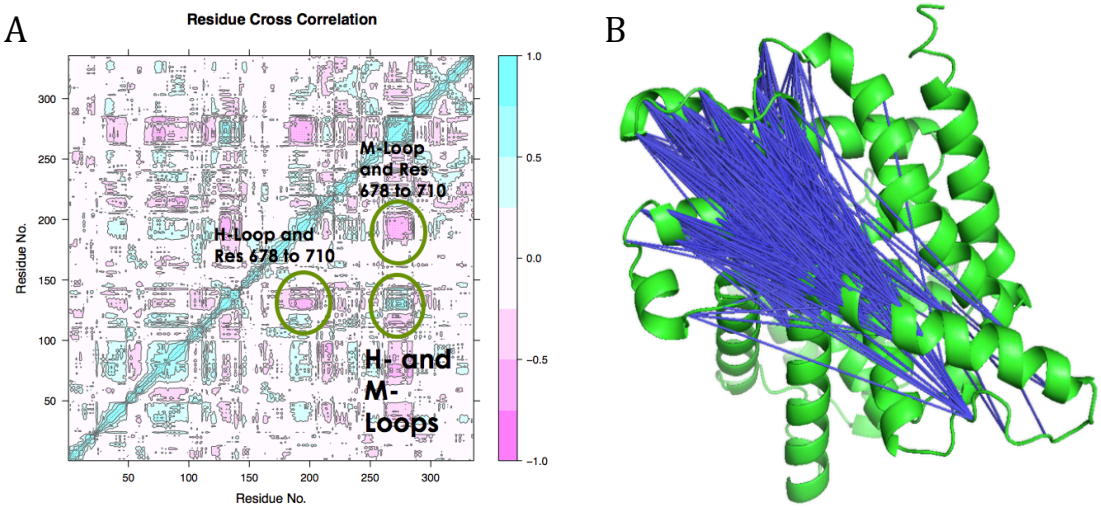


Figure 8: a) Cross-correlational analysis of PDE6 apo system. X and Y axes are residue numbers, and colored regions show correlation (blue) or anti-correlation (purple) for those residues. b) 3D

representation of the anti-correlations between residues in PDE6 apo shown as blue lines that connect the anti-correlated residues.

The Bio3D cross-correlational analysis showed very little correlations for all of the PDE5 and PDE5.6 systems. The system that showed the most correlations was apo PDE6 (Figure 8). The H- and the M-loops show strong correlations. More interestingly, both loops show correlations with the same 'new region' that showed high flexibility. This could be an indication of a potential allosteric region of regulation that has not been previously studied.

Discussion

As mentioned, the validation data was strong enough to be able to confirm that the methods and structures used would give accurate data from stable simulations. The initial aim of this work was to understand in more depth the H- and M-loop distance and dynamics.

The H- and M-loop distances in PDE6 systems being small and stable agrees with the structural information from the crystal structure of PDE5/6, which suggested a smaller H- and M-loop distance than in PDE5. The consistency of this small distance across the sildenafil and γ-bound systems is a good indication that this structural feature is not an artifact of any one system but a robust feature of the protein. The PDE5 H- and M-loop distance varying so wildly may be the structural explanation for the lower rate of cGMP hydrolysis. A small distance may somehow aid in binding of cGMP for hydrolysis, so the stable small distance seen PDE6 yields high efficiency while the varying, unstable distance in PDE5 yields lower efficiency.

The previously noted correlation between the H- and M-loops are not unexpected, the RMSFs showed that the loops were very flexible/mobile and the distance between the loops stayed constant. The only way to reconcile these facts would be if the loops were moving in coordination together, i.e. if there motions of the loops were correlated. The confirmation of this correlation through the cross-correlational analysis neatly ties together all of these data. That these correlations don't exist in the PDE5 or PDE5/6 systems may mean that the two loops behave differently in PDE6. The correlation that exists between the two loops in PDE6 may somehow contribute to increased affinity for cGMP and the resulting higher rate of hydrolysis of cGMP in PDE6.

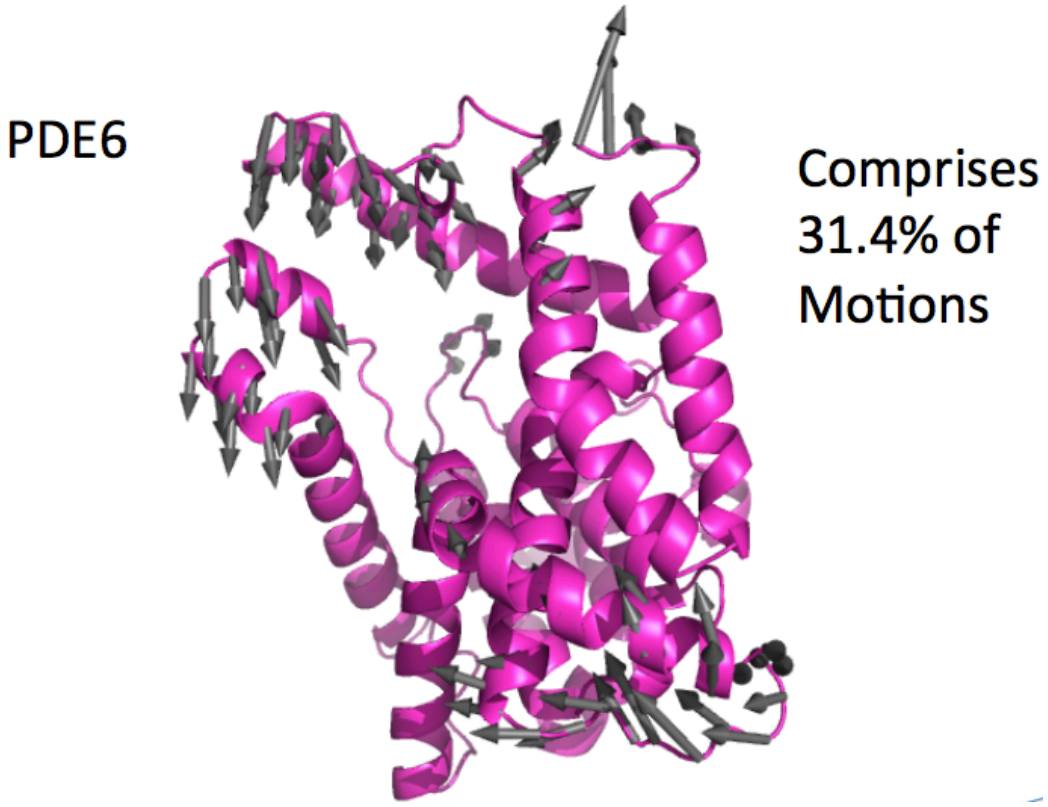


Figure 9: Motions of the first principal component for PDE6 apo.

Some key data pieces may explain the structural reasons for the difference in functionality in PDE6 and PDE5. That the previously mentioned third region (residues

678 to 710, tentatively called the KM region) with high fluctuations seen in the RMSF also showed correlations with both the H- and M-loops indicated that it may be an allosteric region of regulation. The motions of the first principal component (Figure 9), which comprise 31.4% of the motion, also supports this idea. The first principal component shows a 'breathing' motion with the H- and M-loops moving together towards the catalytic pocket and the KM region moving up towards the catalytic pocket. The motions of the first principal component for PDE5 apo do not show this same breathing motion (Supplementary Figure 2). Our hypothesis is that this concerted motion helps provide access and possibly recruit cGMP to the catalytic pocket to allow hydrolysis, which results in the high activity observed in PDE6. The lack of fluctuations and correlations in the corresponding residues in PDE5 and PDE5/6 along with the lack of the breathing motion in the first principal component all has lead us to hypothesize that in these enzymes, cGMP diffuses passively to the catalytic pocket without guidance from the two loops and the KM region.

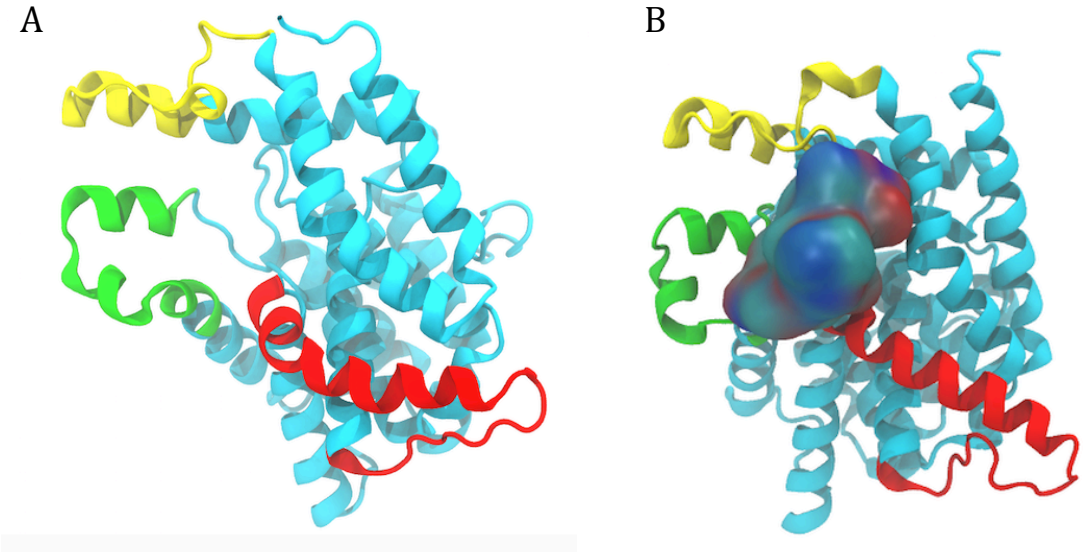


Figure 10: a) PDE6 apo with H-loop highlighted in green, M-loop in yellow, and KM region in red. b) PDE6 with the γ-subunit bound. γ-subunit shown in space filling representation,

highlighting that it blocks the catalytic pocket and interferes with interactions between the H- and M-loops and the KM region.

The KM region may also explain the need for the γ-subunit. As a key regulator of the eyesight signaling pathway, regulation of PDE6 itself is key. When PDE6 is active, it is necessary for the enzyme to hydrolyze cGMP quickly so that the signal can be passed on quickly and vision can occur almost instantly. Our hypothesis is that the KM region works in concert with the H- and M-loops to cause active transport of cGMP to the catalytic pocket for hydrolysis. When there is no light, there should be no signal propagation and no vision, so PDE6 should be completely off. The γ-subunit not only blocks access to the catalytic pocket, it also disrupts any allosteric interactions between the H- and M-loops and the KM region by physically coming between the regions (Figure 10). The blocking of the H- and M-loops and KM region interactions prevents any active diffusion to the catalytic pocket in addition to blocking access to the pocket.

All of the data strongly suggests that the KM region is an allosteric region of regulation. This information can hopefully be used by experimentalists to create a new chimera that more accurately represents PDE6 and its functionality. The KM region may also be used as an allosteric therapeutic target for retinitis pigmentosa or other diseases in which PDE6 is implicated.

The data from these simulations suggests that the current PDE5/6 model may be a poor experimental model of PDE6. As mentioned, the residues that correspond to the KM region did not show high fluctuations or correlations in the chimera. In addition, the entire cross-correlational graph for the chimera showed almost no correlations across all the residues. That the H- and M-loops and the KM region all lacked correlations in the chimera and had strong correlations in PDE6 shows that the two enzymes have very

different structural dynamics. This is further reinforced by the fact that when the γ-subunit was bound to PDE5/6 in the simulations, the fluctuations in the RMSF increased for most of the residues, whereas in PDE6 there was reduced flexibility throughout the protein. All of this in conjunction with the fact that PDE5/6 hydrolyzes cGMP much slower than PDE6 suggests the large difference in the function of the two enzymes is due to these structural differences. By substituting in the residues for the KM region of PDE6 into PDE5, experimentalists may be able to create a chimera that better models PDE6 in its hydrolysis of PDE6 and inhibition by the γ-subunit.

The mutated PDE6 systems showed vary different sampling in the principal component space and had reduced correlations. The two mutations studied are both located on the H-loop on one of the residues closest to the KM region (ASN661), indicated that the mutations may potentially be disrupting the interactions between the H-loop and the KM region and that this disruption of the interactions is what causes such a radical change in the motions of mutated PDE6 in principal component space. It is possible that after further study of the KM region and learning more about its functional significance that it may be used as a therapeutic target to treat retinitis pigmentosa.

Future Directions

Although the data from this simulation study is very promising, there is quite a lot more work to be done in understanding PDE6 and its structure-function relationship. Preliminary work has been done in modeling and simulating a new chimera with the KM region residues from PDE6 substituted in (a gain of function experiment) and a PDE6 model with the KM region residues from PDE5 substituted in (a loss of function experiment). Over the course of 100ns, the gain of function chimera showed increased

correlations and the loss of function PDE6 showed loss of correlations. Although this is promising, more analyses and simulations need to be run to confidently say that substituting in the PDE6 KM region into PDE5/6 creates a chimera that better models PDE6.

More work is also necessary in understanding exactly what occurs in the mutated PDE6 simulations. Do the mutations directly disrupt the H-loop-KM region allosteric interactions or are there long range differences that cause the changes in the motions in principal component space? If there is a direct disruption of interactions, which residues are key to the functional interactions? And how could the KM region potentially be targeted for therapeutic drug design? Along the same lines, performing ligand binding/unbinding simulations could help us further understand how the correlated motions in PDE6 facilitate catalysis. This may help to understand the interactions between the H- and M-loops and the KM region in PDE6 and how they differ from those same interactions in PDE5 and the chimera. Metadynamics is one method that could be used to run simulations in which the ligand (cGMP or sildenafil) is pulled out of the catalytic pockets of PDE5 and PDE6 in an unbiased direction. By observing the pathway that the ligand unbinds, we could evaluate if there are different pathways taken between PDE5 and PDE6 that may relate the catalytic efficiency and degree of cooperativity within the different enzyme. The binding pathway could be 'reverse engineered' and insight could be gained into the structural differences in how the enzymes bind ligands.

Ongoing work is being done to validate the H- and M-loop distances calculated from the 100ns simulations. For each system, 300ns simulations with three different starting velocities are being run. These multiple and longer simulations should give

enough data to validate the H- and M-loop distances and other results reached from all of the 100ns simulations.

Conclusions

The simulations done show that the PDE5/6 H- and M-Loop distance observed experimentally may be indicative of actual PDE6 H- and M-Loop distance, with validation work ongoing (Supplementary Figure 3). We believe that the differences in H-M loop dynamics may relate to different ligand binding pathways and higher catalytic efficiency for PDE6. The KM region (residues 678 to 710) discovered through this work may explain the need for a regulatory subunit (γ) in PDE6, as the PDE6-γ binding location would potentially disrupt the correlations between KM and H-M regions. Finally, PDE5/6 may be an incomplete model for PDE6 because the residues that correspond to the KM region in PDE5/6 did not show high fluctuations in RMSF data or correlations with the H- and M-Loops in dynamic cross correlation data. Using this information, a chimera that more accurately models PDE6 may warrant development.

Acknowledgements

I would like to thank all members of the May Lab for their support, specifically Dr. Eric May for serving as my Honors Advisor and as a great mentor as well as Jason Pattis for his help. I would also like to thank Dr. Mary Bruno and Dr. Victoria Robinson for their help and for serving on my University Scholar committee.

I would also like to acknowledge the computational resources provided through the UConn BECAT center HORNET cluster. Finally, I would like to thank the UConn

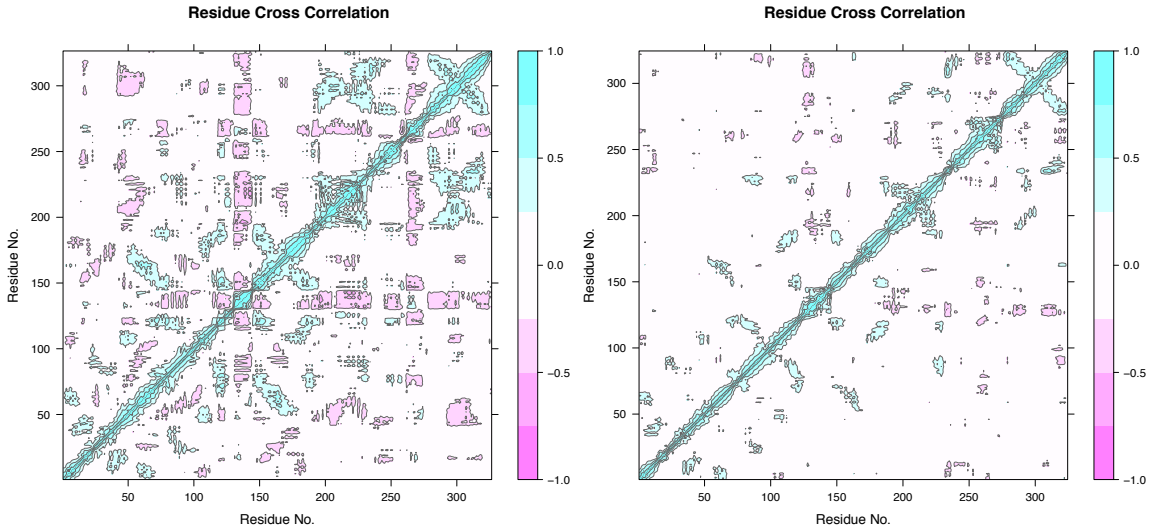
Office of Undergraduate Research for their generous funding through the Honors Life Sciences Thesis Award, the SURF Award, and the IDEA Grant.

References

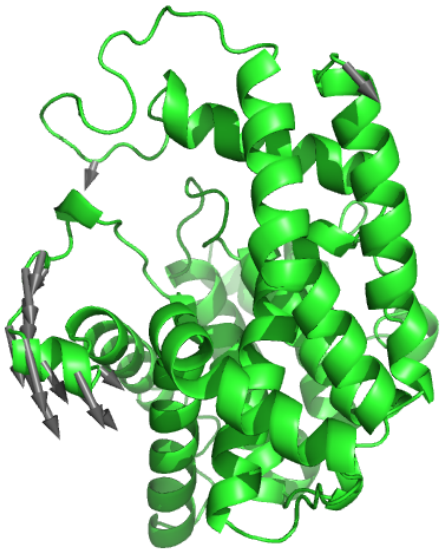
1. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nature Structural & Molecular Biology. Nature Publishing Group; 2002;9(9):646–52.
2. Cote RH. Structure, function, and regulation of photoreceptor phosphodiesterase (PDE6). regulation. 2003;20(30):40.
3. Yang B, Hamza A, Chen G, Wang Y, Zhan C-G. Computational Determination of Binding Structures and Free Energies of Phosphodiesterase-2 with Benzo[1,4]diazepin-2-one Derivatives. J. Phys. Chem. B. 2010 Dec. 9;114(48):16020–8.
4. Zhao P, Chen S-K, Cai Y-H, Lu X, Li Z, Cheng Y-K, et al. The molecular basis for the inhibition of phosphodiesterase-4D by three natural resveratrol analogs. Isolation, molecular docking, molecular dynamics simulations, binding free energy, and bioassay. BBA - Proteins and Proteomics. Elsevier B.V; 2013 Oct. 1;1834(10):2089–96.
5. Goc A, Chami M, Lodowski DT, Bosshart P, Moiseenkova-Bell V, Baehr W, et al. Structural Characterization of the Rod cGMP Phosphodiesterase 6. Journal of Molecular Biology. Elsevier Ltd; 2010 Aug. 20;401(3):363–73.
6. Bender AT. Cyclic Nucleotide Phosphodiesterases: Molecular Regulation to Clinical Use. Pharmacological Reviews. 2006 Sep. 1;58(3):488–520.
7. Song J, Guo L-W, Muradov H, Artemyev NO, Ruoho AE, Markley JL. Intrinsically disordered γ-subunit of cGMP phosphodiesterase encodes functionally relevant transient secondary and tertiary structure. Proceedings of the National Academy of Sciences. National Acad Sciences; 2008;105(5):1505–10.
8. Barren B, Gakhar L, Muradov H, Boyd KK, Ramaswamy S, Artemyev NO. Structural basis of phosphodiesterase 6 inhibition by the C-terminal region of the γ-subunit. The EMBO Journal. Nature Publishing Group; 2009 Oct. 1;28(22):3613–22.
9. Zeng-Elmore X, Gao X-Z, Pellarin R, Schneidman-Duhovny D, Zhang X-J, Kozacka KA, et al. Molecular Architecture of Photoreceptor Phosphodiesterase Elucidated by Chemical Cross-Linking and Integrative Modeling. Journal of Molecular Biology. Elsevier Ltd; 2014 Nov. 11;426(22):3713–28.

10. Arshavsky, Vadim Y., Trevor D. Lamb, and Edward N. Pugh Jr. "G proteins and phototransduction." *Annual review of physiology* 64.1 (2002): 153-187.
11. Striker L. Biochemistry: A Short Course 2nd Edition. 2nd ed. New York: WH Freeman And Company.
12. Arshavsky VY, Burns ME. Photoreceptor Signaling: Supporting Vision across a Wide Range of Light Intensities. Journal of Biological Chemistry. 2012 Jan. 13;287(3):1620–6.
13. Mariani S, Dell'Orco D, Felline A, Raimondi F, Fanelli F. Network and Atomistic Simulations Unveil the Structural Determinants of Mutations Linked to Retinal Diseases. Dunbrack RL, editor. PLoS Comput Biol. 2013 Aug. 29;9(8):e1003207.
14. McCammon, J. Andrew, Bruce R. Gelin, and Martin Karplus. "Dynamics of folded proteins." *Nature* 267.5612 (1977): 585-590.
15. D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen and H. J. C. Berendsen, Gromacs User Manual version 4.5.6, www.gromacs.org (2010)
16. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9(1):40.
17. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. "UCSF Chimera - A Visualization System for Exploratory Research and Analysis." *J. Comput. Chem.* **25**:1605-1612 (2004).
18. Dokholyan NV, editor. Biological and Medical Physics, Biomedical Engineering. Boston, MA: Springer US; 2012.
19. Grant, B.J. et al. (2006) Bioinformatics 22, 2695--2696.
20. Balsera MA, Wriggers W, Oono Y, Schulten K. Principal component analysis and long time protein dynamics. The Journal of Physical Chemistry. ACS Publications; 1996;100(7):2567–72.
21. Wang H, Liu Y, Huai Q, Cai J, Zoraghi R, Francis SH, Corbin JD, Robinson H, Xin Z, Lin G, Ke H (2006) Multiple conformations of phosphodiesterase-5: implications for enzyme function and drug development. J Biol Chem 281: 21469–21479

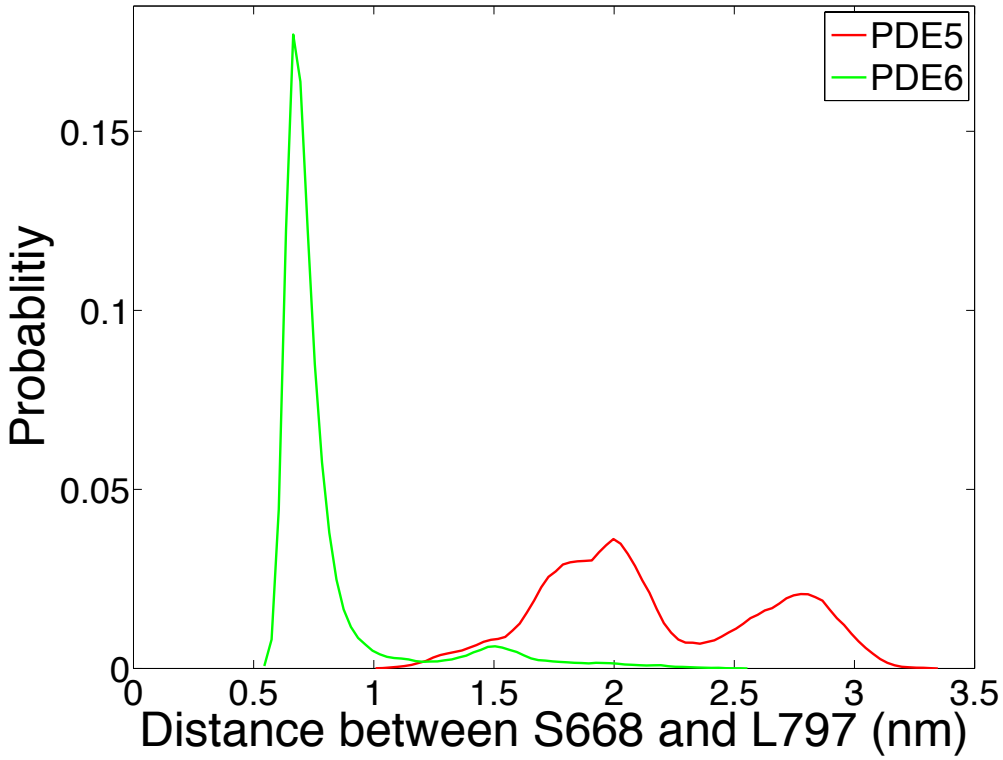
Supplemental Figures



Supplementary Figure 1: A) Cross-correlational analysis of PDE5 apo system. B) Cross-correlational analysis of PDE5/6 apo system



Supplementary Figure 2: Motions of the first principal component of PDE5 apo. Comprises 32.7% of the motions of the protein during the simulation.



Supplementary Figure 3: Histogram of PDE5 and PDE6 apo H- and M-loop distances after preliminary 300ns simulation of each system.